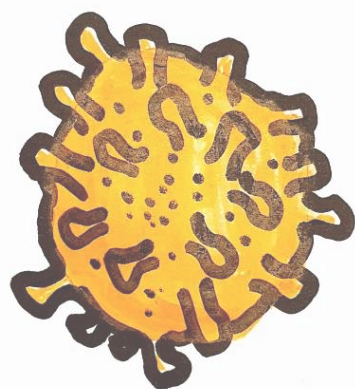


# ICEI Papers COVID-19

*Instituto Complutense de Estudios Internacionales*

---



**Nº 12**

**6 DE ABRIL DE 2020**

---

**UNA BREVE INTRODUCCIÓN AL  
MODELO SIR APLICADO AL CASO  
DEL COVID-19**

Víctor Martín Barroso



# Una breve introducción al modelo SIR aplicado al caso del Covid-19

Víctor Martín Barroso

*Investigador adscrito al Instituto Complutense de Estudios Internacionales (ICEI) y Profesor Titular de Economía Aplicada en la Universidad Rey Juan Carlos (URJC).*

## EL MODELO SIR

Dentro de la epidemiología, uno de los modelos matemáticos determinísticos más usado es el denominado modelo SIR (*Susceptible-Infectious-Recovered*). El modelo, que tiene su origen en el trabajo de Kermack y McKendrick (1972)<sup>1</sup>, permite caracterizar la evolución de una epidemia que se propaga mediante contagio entre los individuos de una población de tamaño  $N$  que permanece constante. En cada momento del tiempo  $t$  los individuos pueden encontrarse en tres posibles estados o compartimentos,

- **Susceptibles  $S(t)$ :** individuos sanos que pueden ser contagiados por los infectados.
- **Infectados  $I(t)$ :** individuos portadores de la enfermedad.
- **Recuperados  $R(t)$ :** individuos que se han recuperado de la enfermedad<sup>2</sup>.

El modelo cuenta con dos parámetros, que son los siguientes:

- **Tasa de infección o tasa de transmisión de la enfermedad  $\beta$ :** nos indica el número de contactos suficientes para la transmisión de una persona por unidad de tiempo. Es decir, el parámetro  $\beta$  es el producto del número de contactos medios por individuo y unidad de tiempo y la probabilidad de contagio en un solo contacto entre un individuo infectado y un individuo susceptible. En consecuencia  $\beta I(t)/N$  nos indica el número medio de contactos infectados por unidad de tiempo de un individuo susceptible, mientras que  $\beta S(t)I(t)/N$  es el número de casos nuevos por unidad de tiempo.
- **Tasa de recuperación  $\gamma$ :** nos indica la tasa por unidad de tiempo a la cual los individuos infectados se recuperan. Es decir, es la ratio entre el número de recuperados y el número de infectados por unidad de tiempo. El periodo medio de latencia se calcula como la inversa de la tasa de recuperación,  $L = 1/\gamma$ .

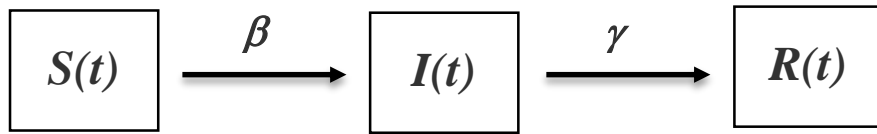
Es importante resaltar, de nuevo, que el valor de la tasa de infección o transmisión está relacionado, por un lado, con el número de contactos por unidad de tiempo y, por otro lado, con la probabilidad de contagio. Si bien es cierto que una parte importante de la tasa de infección depende de si la enfermedad es más o menos contagiosa, las medidas de prevención como el confinamiento, el distanciamiento y la adopción de medidas de higiene y protección por parte de la población, tienden a reducir el valor de la misma.

---

<sup>1</sup> Kermack, W.O. y McKendrick (1972). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*. Vol. 115, No. 772, pp. 700-721.

<sup>2</sup> El conjunto de recuperados lo conforman tanto los individuos que se han recuperado de la enfermedad y se han vuelto inmunes, como por el conjunto de individuos que fallecen a consecuencia de la enfermedad.

El esquema correspondiente a este modelo matemático es el siguiente,



La dinámica del modelo está representada por el siguiente sistema de ecuaciones diferenciales que describen la variación en el tiempo de cada uno de los tres estados posibles,

$$\frac{dS(t)}{dt} = -\beta S(t) \frac{I(t)}{N} \quad (1)$$

$$\frac{dI(t)}{dt} = \beta S(t) \frac{I(t)}{N} - \gamma I(t) \quad (2)$$

$$\frac{dR(t)}{dt} = \gamma I(t) \quad (3)$$

Partiendo de un número inicial dado de individuos infectados ( $I_0$ ), las condiciones iniciales del modelo son las siguientes,

$$S(0) = S_0 = N - I_0, \quad I(0) = I_0, \quad R(0) = 0 \quad (4)$$

Un parámetro de gran interés en epidemiología es la denominada tasa básica de reproducción<sup>3</sup>  $R_0$ , utilizada para describir la intensidad de una enfermedad infecciosa. De forma específica, la tasa reproductiva básica nos indica el número de nuevos infectados, en promedio, por un solo infectado a lo largo del periodo de duración de la enfermedad infecciosa. La tasa de reproducción básica se calcula de la siguiente manera:

$$R_0 = \frac{\beta}{\gamma} \quad (5)$$

La tasa reproductiva básica nos da indicios de si la enfermedad se va a propagar o no, y de la velocidad con que una enfermedad puede propagarse en una población de individuos determinada. Cuando  $R_0 > 1$  la enfermedad puede llegar a propagarse ampliamente entre la población. Sin embargo, si  $R_0 < 1$  la enfermedad desaparece tras un cierto periodo de tiempo.

Este sencillo modelo, que permite una gran casuística mediante diversidad de variantes del modelo básico, presenta al menos dos utilidades para el caso que nos concierne. En primer lugar, permite simular la evolución de una determinada enfermedad infecciosa y comparar distintas situaciones o escenarios, añadiendo al modelo valores para los siguientes parámetros: tasa de infección ( $\beta$ ), tasa de recuperación ( $\gamma$ ), tamaño de la población ( $N$ ), y número de infectados inicialmente ( $I_0$ ). En segundo lugar, es posible ajustar el modelo SIR a datos reales procedentes de un determinado país o región, lo que permite obtener estimaciones sobre la tasa de infección, la tasa de recuperación y, en

---

<sup>3</sup> La primera aplicación moderna de la tasa básica de reproducción se encuentra en MacDonald, G. (1952). *The Epidemiology and control of Malaria*. London: Oxford University Press.

consecuencia, sobre la tasa reproductiva básica. De este modo se obtiene información valiosa de cómo está evolucionando la enfermedad.

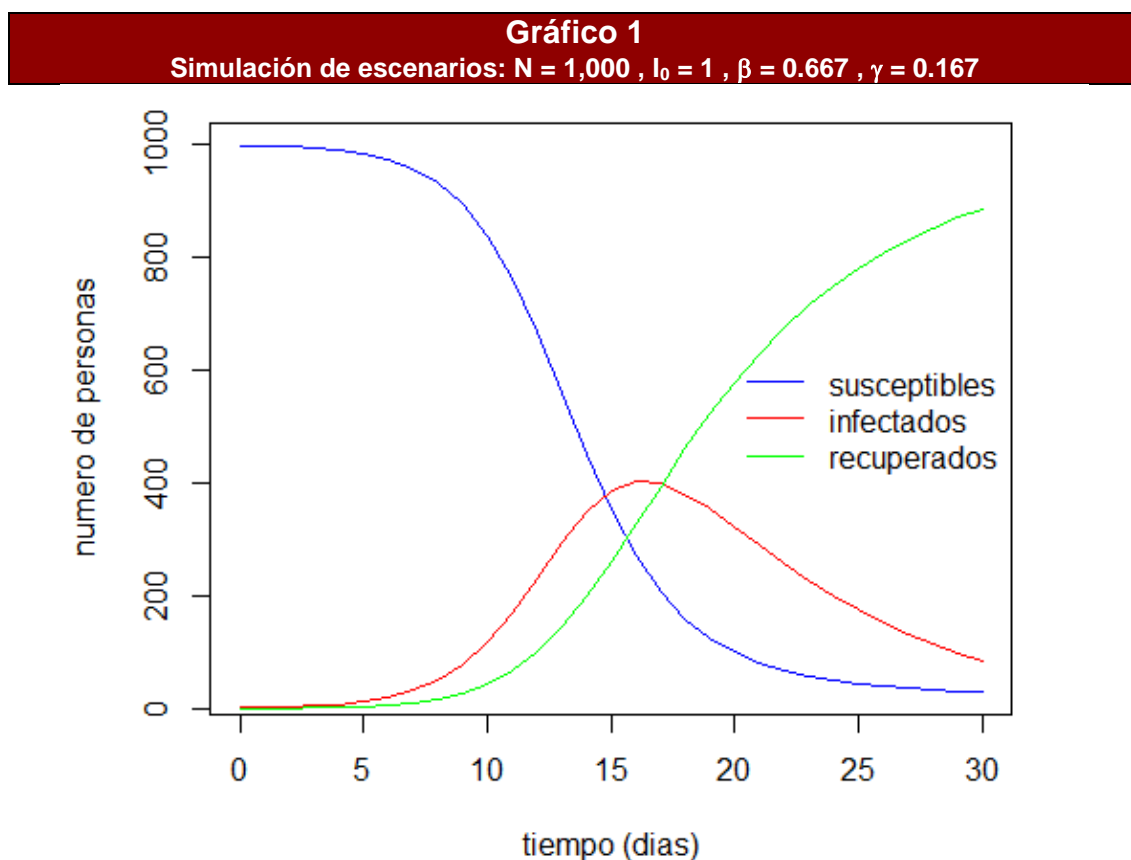
## SIMULACIÓN DE ESCENARIOS

Mediante el uso del paquete “deSolve”<sup>4</sup> en R, es posible simular y comparar la evolución de una epidemia siguiendo las ecuaciones dadas por el modelo SIR bajo distintos escenarios.

A modo de ejemplo, se ha simulado la evolución de una epidemia en una región con una población de 1,000 individuos, y en donde el número inicial de infectados es de una persona. Asimismo, se ha supuesto una tasa de transmisión de 0.667 y una tasa de recuperación de 0.167, lo que equivale a un periodo de latencia de 6 días. Estos datos implican una tasa básica de reproducción de 4.

El gráfico 1 muestra la evolución durante un periodo de 30 días en el número de individuos susceptibles, infectados y recuperados. Como puede apreciarse, el pico de infectados se produce poco después de dos semanas desde el inicio de la epidemia (día 17), en el cuál se registran un total de 402 nuevos infectados. Justo en el pico de la epidemia, el número total de individuos infectados o enfermos asciende a 947 individuos, lo que representa el 94,7% de la población.

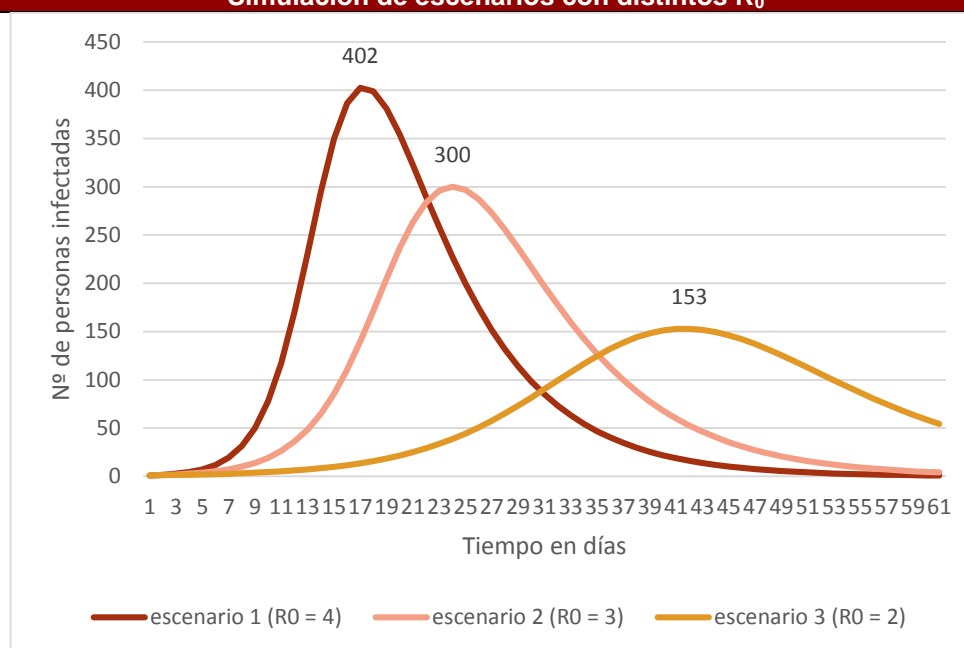
Por su parte, el gráfico 2 presenta la evolución en el número de personas infectadas en tres posibles escenarios distintos, en donde las diferencias se producen exclusivamente en la tasa de transmisión de la enfermedad, y por tanto en la tasa reproductiva básica.



<sup>4</sup> El paquete deSolve permite resolver sistema de ecuaciones diferenciales. Más información disponible en: <https://cran.r-project.org/web/packages/deSolve/index.html>



**Gráfico 2**  
**Simulación de escenarios con distintos  $R_0$**



Fuente: elaboración propia mediante el programa estadístico R.

Los valores asignados a los distintos parámetros del modelo SIR en los tres escenarios son los siguientes:

- Escenario 1:  $N = 1,000$  ,  $I_0 = 1$  ,  $\beta = 0.667$  ,  $\gamma = 0.167 \rightarrow R_0 = 4$
- Escenario 2:  $N = 1,000$  ,  $I_0 = 1$  ,  $\beta = 0.500$  ,  $\gamma = 0.167 \rightarrow R_0 = 3$
- Escenario 3:  $N = 1,000$  ,  $I_0 = 1$  ,  $\beta = 0.333$  ,  $\gamma = 0.167 \rightarrow R_0 = 2$

El gráfico muestra que, para tasas de infección o transmisión menores, el pico de la infección se retrasa en el tiempo, y se reduce el número máximo de personas infectadas por día de forma considerable. Para una tasa de infección de 0.5 (escenario 2) el pico de la epidemia tiene lugar 24 días después del inicio de la epidemia, con un total de 300 nuevos infectados, y un total de 421 personas infectadas o enfermas (un 42,1% de la población). Para una tasa de infección de 0.333 (escenario 3) el pico de la epidemia tiene lugar 42 días después del inicio de la epidemia (casi dos veces el tiempo que en el escenario 1) con un total de 153 nuevos infectados.

## APLICACIÓN AL CASO DEL COVID-19 EN ESPAÑA

Con el fin de ilustrar el ajuste de un modelo SIR a un caso real, se han utilizado los datos diarios disponibles sobre el número de infectados por el covid-19 en España, procedentes del Centro Europeo para el Control y Prevención de Enfermedades<sup>5</sup> (ECDC). En el momento de la extracción, la serie de datos está disponible desde el 04-02-2020 hasta el 05-04-2020. El proceso de obtención de estimaciones para los parámetros  $\beta$  y  $\gamma$ , requiere la minimización de la suma de cuadrados de los

<sup>5</sup> Los datos están disponibles en <https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide>

residuos (RSS), siendo los residuos la diferencia entre el número real de personas infectadas y el número estimado de personas infectadas a partir del modelo SIR, es decir,

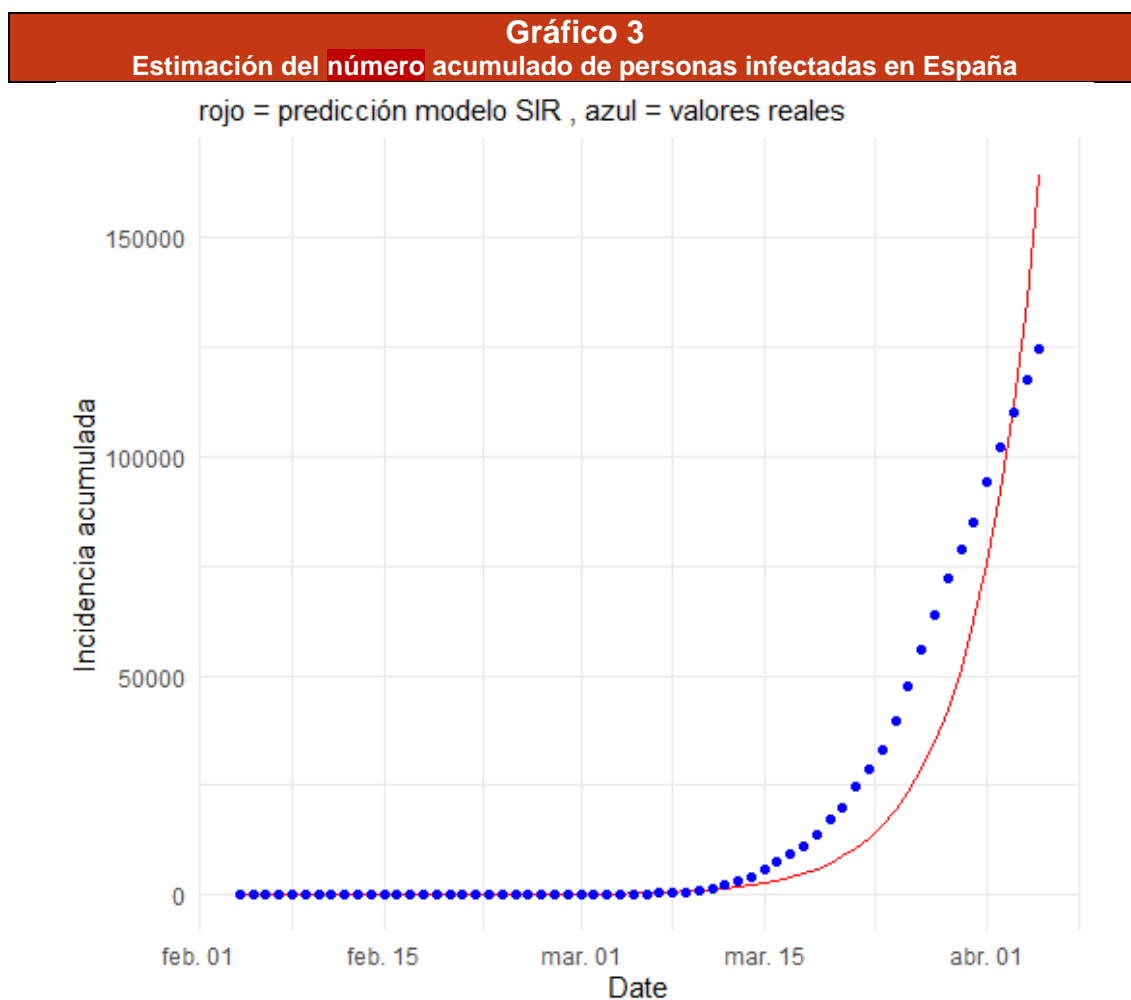
$$RSS = \sum_i \left( I(t) - \hat{I}(t) \right)^2 \quad (6)$$

Los resultados obtenidos mediante el programa deSolve en R son los siguientes,

$$\hat{\beta} = 0.5987 \quad , \quad \hat{\gamma} = 0.4013 \quad , \quad \hat{R}_0 = 1.49 \quad (7)$$

La tasa de infección es prácticamente igual a 0.6. La tasa de recuperación es de 0.4 lo que implica un periodo de latencia de 2.5 días. La tasa básica de reproducción se sitúa en 1.5, lo que significa que, en España, en promedio, cada persona infectada es capaz de infectar a 1.5 personas.

El gráfico 3 muestra el ajuste del modelo estimado. Como puede apreciarse, el modelo predice un crecimiento de la enfermedad más exponencial que el realmente observado. Esta diferencia podría deberse a las distintas medidas de confinamiento y limitaciones de la actividad económica que se han ido produciendo en España desde el 11 de marzo, así como las medidas de higiene y protección llevadas a cabo por la población pues, en conjunto, todo ello supone una reducción de la tasa de transmisión del covid-19 tanto por la disminución en el número de contactos diarios por persona y día, como por la reducción en la tasa de contagio.



Fuente: elaboración propia mediante el programa estadístico R.

## ANEXO

### Código R para reproducir el gráfico 1

```
# Cargamos paquete deSolve
library(deSolve)
# Indicamos las ecuaciones diferenciales del modelo SIR
sir_equations <- function(time, variables, parameters) {
  with(as.list(c(variables, parameters)), {
    dS <- -beta * I * S / N
    dI <- beta * I * S / N - gamma * I
    dR <- gamma * I
    return(list(c(dS, dI, dR)))
  })
}

# Indicamos parámetros iniciales
N <- 1000
parameters_values <- c(
  beta = 0.333,
  gamma = 0.167
)
initial_values <- c(
  S = N - 1,
  I = 1,
  R = 0
)

# Indicamos el nº de días a simular
time_values <- seq(0, 60)

# Simulamos
sir_values_1 <- ode(
  y = initial_values,
  times = time_values,
  func = sir_equations,
  parms = parameters_values
)

# Creamos el gráfico 1
sir_values_1 <- as.data.frame(sir_values_1)
with(sir_values_1, {
  plot(time, S, type = "l", col = "blue",
    xlab = "tiempo (días)", ylab = "numero de personas")
  lines(time, I, col = "red")
  lines(time, R, col = "green")
})
legend("right", c("susceptibles", "infectados", "recuperados"),
  col = c("blue", "red", "green"), lty = 1, bty = "n")
```

### Código R para ajustar el modelo SIR a los datos de España

```
# Indicamos las ecuaciones diferenciales del modelo SIR
SIR <- function(time, state, parameters) {
  par <- as.list(c(state, parameters))
  with(par, {
    dS <- -beta * I * S / N
    dI <- beta * I * S / N - gamma * I
    dR <- gamma * I
    list(c(dS, dI, dR))
  })
}

# Cargamos los datos del ECDC, seleccionamos España y calculamos la incidencia acumulada
library(readxl)
library(httr)
```

```

url <- paste("https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-geographic-disbtribution-worldwide-",
",format(Sys.time(), "%Y-%m-%d"), ".xlsx", sep = "")
GET(url, authenticate(":", ":", type="ntlm"), write_disk(tf <- tempfile(fileext = ".xlsx")))
data <- read_excel(tf)
df <- data %>%
  dplyr::filter(countriesAndTerritories == "Spain") %>%
  dplyr::arrange(dateRep) %>%
  dplyr::mutate(
    confirmed_cum = cumsum(cases),
    death_cum = cumsum(deaths)
  )
library(lubridate)
Infected <- subset(df, dateRep >= ymd("2020-02-04") & dateRep <= ymd("2020-04-05"))$confirmed_cum
Day <- 1:(length(Infected))

#Indicamos parámetros iniciales, población España en 2019 46.7 millones
N <- 46660000
init <- c(
  S = N - Infected[1],
  I = Infected[1],
  R = 0
)

# Minimizamos la suma de residuos al cuadrado, indicamos los valores iniciales de beta y gamma
RSS <- function(parameters) {
  names(parameters) <- c("beta", "gamma")
  out <- ode(y = init, times = Day, func = SIR, parms = parameters)
  fit <- out[, 3]
  sum((Infected - fit)^2)
}
library(deSolve)
Opt <- optim(c(0.5, 0.5),
  RSS,
  method = "L-BFGS-B",
  lower = c(0, 0),
  upper = c(1, 1)
)
Opt$message
Opt_par <- setNames(Opt$par, c("beta", "gamma"))
Opt_par

# Generamos los valores predcidos por el modelo
sir_start_date <- "2020-02-04"
t <- 1:as.integer(ymd("2020-04-6") - ymd(sir_start_date))
fitted_cumulative_incidence <- data.frame(ode(
  y = init, times = t,
  func = SIR, parms = Opt_par
))

# Generamos el gráfico 3
library(dplyr)
fitted_cumulative_incidence <- fitted_cumulative_incidence %>%
  mutate(
    Date = ymd(sir_start_date) + days(t - 1),
    Country = "Spain",
    cumulative_incident_cases = Infected
  )
library(ggplot2)
fitted_cumulative_incidence %>%
  ggplot(aes(x = Date)) +
  geom_line(aes(y = I), colour = "red") +
  geom_point(aes(y = cumulative_incident_cases), colour = "blue") +
  labs(
    y = "Incidencia acumulada",
    subtitle = "rojo = predicción modelo SIR , azul = valores reales"
  ) +
  theme_minimal()

```